GENETIC AND EPIGENETIC STUDIES OF DIABETES AND DIABETIC NEPHROPATHY WITH FOCUS ON THE IGF-IGFBP AXIS

Tianwei Gu

Stockholm 2014
To my dearest parents
献给我亲爱的父母
LIST OF PUBLICATIONS


Other publications not included in this thesis:


Ma J, Abu Seman N, Gu T, Ren Z, Zhao L, Lin D, Brismar K, Gu HF. Genotypic prediction of the intercellular adhesion molecule 1 (ICAM1) K469E polymorphism in diabetes and diabetic nephropathy. (Submitted manuscript).
CONTENTS

1 Background .................................................................................................................. 1
  1.1 Diabetes and diabetic nephropathy ................................................................. 1
    1.1.1 Type 1 diabetes ....................................................................................... 2
    1.1.2 Type 2 diabetes ....................................................................................... 2
    1.1.3 Obesity and body mass index .................................................................. 3
    1.1.4 Diabetic nephropathy .............................................................................. 3
    1.1.5 Heterogeneity of diabetes and diabetic nephropathy ....................... 4
  1.2 Genetic studies in diabetes and diabetic nephropathy ..................................... 4
    1.2.1 Heritability in diabetes and diabetic nephropathy .............................. 4
    1.2.2 Human genome and genetic study approaches ...................................... 4
    1.2.3 Genetic studies in diabetes ................................................................. 5
    1.2.4 Genetic studies in diabetic nephropathy .............................................. 6
  1.3 Epigenetic studies in diabetes and diabetic nephropathy ............................ 7
    1.3.1 The basics of epigenetics ........................................................................ 7
    1.3.2 The basics of DNA methylation ........................................................... 7
    1.3.3 The strategies of DNA methylation analysis ......................................... 9
    1.3.4 DNA methylation studies in diabetes and diabetic nephropathy ....... 9
  1.4 IGF-IGFBP axis .................................................................................................. 10
    1.4.1 Overview of IGF-IGFBP axis ............................................................... 10
    1.4.2 IGF-IGFBP axis in glucose homeostasis and diabetes ..................... 13
    1.4.3 IGF-IGFBP axis in diabetic nephropathy ............................................ 14
    1.4.4 Genetic studies in IGF-IGFBP axis ...................................................... 15
  2 Aims ......................................................................................................................... 17
    2.1 General hypothesis ....................................................................................... 17
    2.2 Aims ............................................................................................................. 17
  3 Materials and methods ......................................................................................... 18
    3.1 Subjects ........................................................................................................ 18
      3.1.1 Stockholm diabetes prevention program (SDPP) .............................. 18
      3.1.2 Genetics of kidneys in diabetes study (GoKinD) ............................. 18
      3.1.3 Subjects from Czech Republic .......................................................... 19
      3.1.4 Swedish subjects with type 1 diabetes and diabetic nephropathy .... 19
    3.2 Animal model ............................................................................................... 19
      3.2.1 db/db mice .......................................................................................... 19
    3.3 Methods ......................................................................................................... 20
      3.3.1 Taqman allelic discrimination .............................................................. 20
      3.3.2 Bisulfite pyrosequencing DNA methylation analysis ........................ 20
      3.3.3 Serum IGF-I and IGFBP-1 measurements ......................................... 22
      3.3.4 Insulin and HbA1c measurements ..................................................... 22
      3.3.5 Analyses of mRNA expression and protein levels ............................ 22
    3.4 Statistics ......................................................................................................... 22
  4 Results .................................................................................................................... 24
    4.1 Study I .......................................................................................................... 24
    4.2 Study II & III ................................................................................................ 24
    4.3 Study IV ........................................................................................................ 26
  5 Discussion and future perspective ...................................................................... 27
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td><em>IGF1</em> SNP and DNA methylation</td>
<td>27</td>
</tr>
<tr>
<td>5.2</td>
<td><em>IGFBP1</em> DNA methylation alteration and insulin</td>
<td>27</td>
</tr>
<tr>
<td>5.3</td>
<td>Circulating IGFBP-1 levels in T1D, T2D and DN</td>
<td>28</td>
</tr>
<tr>
<td>5.4</td>
<td>DNA methylation analyses with blood and tissue samples</td>
<td>28</td>
</tr>
<tr>
<td>5.5</td>
<td>Genetic association of <em>IGF2BP2</em> with T2D and T1D-DN</td>
<td>29</td>
</tr>
<tr>
<td>5.6</td>
<td>Other study related to epigenetic mechanism</td>
<td>30</td>
</tr>
<tr>
<td>6</td>
<td>Conclusions</td>
<td>31</td>
</tr>
<tr>
<td>7</td>
<td>Acknowledgements</td>
<td>32</td>
</tr>
<tr>
<td>8</td>
<td>References</td>
<td>34</td>
</tr>
</tbody>
</table>
LIST OF ABBREVIATIONS

ACEI  Angiotensin-converting-enzyme inhibitor
ALS    Acid-labile subunit
ARB    Angiotensin-II receptor blocker
BMI    Body mass index
CNV    Copy number variations
CKD    Chronic kidney disease
DCCT   Diabetes control and complication trial
DN     Diabetic nephropathy
DNMT   DNA-methyltransferase
EWAS   Epigenome-wide association study
ESRD   End-stage renal disease
FHD    Family history of diabetes
GDM    Gestational diabetes mellitus
GH     Growth hormone
GMMDR  Generalized multifactor dimensionality reduction
GoKinD Genetics of kidneys in diabetes study
GK     Goto-Kakizaki
GWAS   Genome-wide association study
HATs   Histone acetyltransferases
HbA1c  Glycated hemoglobin
HDACs  Histone deacetylases
HOMA   Homeostatic model assessment
IFG    Impaired fasting glucose
IGF    Insulin-like growth factor
IGFBP  Insulin-like growth factor binding protein
IGF2BP2 Insulin-like growth factor 2 mRNA binding protein 2
IGT    Impaired glucose tolerance
IR     Insulin resistance
KCNJ11 Potassium inwardly-rectifying channel, subfamily J, member 11
MBD    Methyl-CpG binding protein
miRNA  MicroRNAs
MODY   Maturity-onset diabetes of the young
NGSP   National glycohemoglobin standardization program
NGT    Normal glucose tolerance
OGTT   Oral glucose tolerant test
PPARG  Peroxisome proliferator-activated receptor gamma
SDPP   Stockholm diabetes prevention program
SNP    Single nucleotide polymorphism
TCF7L2  Transcription factor 7-like 2
UTR    Untranslated region
VNTR   Variable number tandem repeats
WC     Waist circumference
WHR    Waist-to-hip ratio
1 BACKGROUND

1.1 DIABETES AND DIABETIC NEPHROPATHY

Diabetes mellitus is a metabolic disease characterized by elevated blood glucose levels resulting from inadequate insulin secretion with or without insulin resistance. The common symptoms with diabetes include increased thirst, increased hunger, frequent urination and weight loss. If untreated or poorly controlled, diabetes can lead to acute life-threatening complications such as nonketotic hyperosmolar syndrome and diabetic ketoacidosis, or progress to chronic complications affecting the eye, kidney, heart, nervous system and other organs. Diabetes is a growing chronic disease pandemic and one of the greatest threats to public health. The global prevalence of diabetes was 382 million in 2013, and is expected to reach 592 million by 2035 [1]. The trend of increasing appears to be more intense among developing countries particularly in China and India. In China, the prevalence of diabetes was reported to be 5.5% in 2000, and this number has doubled reaching to 11.6% in 2013 [2].

The diagnosis of diabetes is primarily based on the levels of plasma glucose, either fasting plasma glucose or 2h plasma glucose after a 75g oral glucose tolerant test (OGTT) [3]. Glycated hemoglobin (HbA1c) is an indicator of average blood glucose level for approximately the past 120 days. Using a method certified by National Glycohemoglobin Standardization Program (NGSP) and standardized or traceable to the Diabetes Control and Complication Trial (DCCT) reference assay, HbA1c ≥6.5% has been added into diabetes diagnosis criteria since the year of 2009 [4]. Pre-diabetes, which is also called intermediate hyperglycemia, refers to individuals with impaired fasting glucose (IFG) and/or impaired glucose tolerance (IGT). The individuals with pre-diabetes have increased risk for the future development of diabetes. It was estimated that up to 70% of the subjects with pre-diabetes will eventually develop diabetes [5]–[8]. The diagnose criteria of diabetes and pre-diabetes [3], [9] are listed in Table 1.

Table 1. Diagnosis of diabetes and pre-diabetes.

<table>
<thead>
<tr>
<th></th>
<th>Diabetes</th>
<th>Pre-diabetes</th>
</tr>
</thead>
<tbody>
<tr>
<td>HbA1c</td>
<td>≥ 6.5%</td>
<td>5.7-6.4%</td>
</tr>
<tr>
<td>Fasting plasma glucose</td>
<td>or ≥ 7.0 mmol/L</td>
<td>or 6.1–6.9 mmol/L (IFG) (WHO) 5.6–6.9 mmol/L (IFG) (ADA)</td>
</tr>
<tr>
<td>OGTT 2h-postload glucose</td>
<td>or ≥ 11.1 mmol/L</td>
<td>or 7.8-11.0 mmol/L (IGT)</td>
</tr>
</tbody>
</table>

WHO: World Health Organization; ADA: American Diabetes Association

Diabetes is classified into type 1 diabetes (T1D), type 2 diabetes (T2D), gestational diabetes mellitus (GDM) and other types of diabetes. T1D and T2D are most common types of diabetes. GDM is diabetes diagnosed during pregnancy that may remit after delivery[3]. Other types of diabetes are diabetes with less common causes, e. g., genetic defects β cell function or drug-induced diabetes. For example, maturity-onset diabetes of the young (MODY) is a form of early-onset, non-insulin dependent diabetes. It compromises a group of monogenetic disorders with impaired β cell function and takes up approximately 1% of all forms of diabetes.
Long-term hyperglycemia in diabetes leads to late chronic complications which are classified based on types of blood vessel involved (Figure 1). These include: i) macro-vascular complications (cerebrovascular disease, cardiovascular disease and peripheral vascular disease); ii) micro-vascular complications (retinopathy, nephropathy and neuropathy) [10].

*Figure 1. Chronic diabetic complications*

![Chronic diabetic complications diagram](image)

*Modified from Bate and Jerums, 2003 [10]*

1.1.1 **Type 1 diabetes**
T1D accounts for 5-10% of all diabetes cases and results from immune-mediated destruction of insulin-secreting pancreatic β cells. T1D is most commonly occurs in children, but it can be diagnosed at any age [11]. The incidence of T1D is increasing globally despite the incidence varies considerably among different countries, from the lowest 0.1/100,000 in China to highest over 60/100,000 in Finland. Sweden has the second highest incidence of T1D in children [11], [12]. More than 90% newly diagnosed T1D patients have one or more autoantibodies against β cell antigens, which is the key feature to distinguish T1D and T2D [11]. Lifetime treatment with exogenous insulin is required for the patients with T1D [13].

1.1.2 **Type 2 diabetes**
T2D accounts for the majority of diabetes cases (approximately 90%). It is characterized by hyperglycemia caused by inadequate pancreatic β cells insulin secretion with or without insulin resistance. Several risk factors are associated with T2D, including obesity, physical inactivity, family history of diabetes (FHD) and genetic defects. The prevalence of T2D is increasing rapidly over the world. Developing countries have more rapidly rising incidence, due to the combined effects
of population growth, urbanization, excessive dietary energy, aging and stress. The prevalence of T2D grows in all age groups including children and adolescents, although T2D is mostly diagnosed in adulthood [14], [15]. Insulin resistance (IR), a hallmark of T2D pathophysiology, stands for impaired insulin action in peripheral tissues (liver, muscle and adipose tissues), and it is usually accompanied by increased insulin production from pancreatic β cells. T2D occurs when β cell function is not able to adapt to the change of insulin action [16]. The method of homeostatic model assessment (HOMA) is usually used to quantify IR and β cell function from basal glucose and insulin concentrations.

1.1.3 Obesity and body mass index

Obesity is defined as the accumulation of excess body fat. Obesity is associated with a number of metabolic disorders including IR, T2D, dyslipidemia, hypertension and cardiovascular diseases. In population studies, body mass index (BMI) is the most commonly-used indicator to evaluate the degree of obesity. In Caucasian adults, individuals with BMI between 25.0-29.9 kg/m² are defined as overweight, while those have BMI over 30 kg/m² are obese [17]. Some other proxy-markers are also used for the assessment of obesity, such as waist circumference (WC) and waist-to-hip ratio (WHR), which are more specific to reflect abdominal obesity [18].

1.1.4 Diabetic nephropathy

Diabetic nephropathy (DN) is the leading cause of chronic kidney disease (CKD) in patients with renal transplantation therapy and it occurs in 20-40% patients with diabetes [19]. The racial and ethnic differences in DN prevalence have been clearly demonstrated by epidemiological studies. Compared to Caucasian whites, the prevalence of diabetic renal disease is higher in African Americans, Hispanics, Asians and Native Americans [20], [21]. Patients with DN exhibit persistent proteinuria, declined renal function, hypertension and increased susceptibility to cardiovascular disease. DN is characterized by progressive pathophysiological changes in kidney, beginning with glomerular hyper-filtration, glomerular hypertrophy, subsequent reduction of glomerular filtration rate (GFR) and eventually progressing to overt nephropathy [22].

DN progresses gradually over the years. The stage of CKD is based on the values of GFR, whereas DN is categorized into stages according to the degree of proteinuria. Persistent proteinuria in the range of 30-299 mg/24h, which is also termed as microalbuminuria, is an early stage of DN. Persistent proteinuria over 300 mg/24h (also called macroalbuminuria) are usually considered to be overt nephropathy [3]. It is estimated that 30-40% diabetic patients with microalbuminuria will develop overt nephropathy over 5-10 years [23]. Patients with microalbuminuria or macroalbuminuria are usually treated with medication such as angiotensin-converting-enzyme inhibitor (ACEI) or angiotensin-II receptor blocker (ARB). CKD can progress to end stage renal disease (ESRD), which is also known as the fifth stage of CKD. The patients with ESRD have a GFR less than 15 ml/min/1.73 m² body surface areas and require dialysis or renal transplantation.
There are similarities and differences between T1D- and T2D-DN. As an important determinant of nephropathic status, chronic uncontrolled hyperglycemia is closely associated with worsening renal function in both T1D and T2D. In T1D, most patients have elevated blood glucose at young ages and have longer duration of diabetes before onset of DN. In contrast, the onset ages of T2D patients are usually much older with shorter duration before development of DN. Besides of hyperglycemia, the renal function in T2D patients may have been affected by the age and other chronic renal injury promoters, such as hypertension, hyperlipidemia, obesity and smoking [24].

1.1.5 Heterogeneity of diabetes and diabetic nephropathy

Diabetes and DN are heterogeneous diseases resulting from the combined effects of genetic inheritance, epigenetic regulation and environmental factors. All these factors not only work independently, but also interact with each other in the pathogenesis of diabetes and its complications.

1.2 GENETIC STUDIES IN DIABETES AND DIABETIC NEPHROPATHY

1.2.1 Heritability in diabetes and diabetic nephropathy

Genetic inheritance is an important predisposing factor in the development of both diabetes and DN. Familiar clustering of diabetes has been shown by the high concordance rates in monozygotic twins and increased risk in individuals with affected first-degree relative [25]–[27]. Family aggregation of DN has also been demonstrated. The fact that some patients with poorly controlled blood glucose do not develop DN, whereas many patients develop DN within a relatively short time after diagnosis of diabetes despite of tight glycemic control, implies the glycemic control and diabetic duration cannot fully explain the development of nephropathy [28]. Nephropathy in the proband is considered to be the only significant predictor of the renal status in the diabetic siblings [29]. Diabetic siblings of the proband with nephropathy was reported to have 2.3 times higher risk of DN compared with siblings of proband free of DN [30].

The above evidence supported that the genetic components are contributed in the development of diabetes and DN.

1.2.2 Human genome and genetic study approaches

Human genome is a complete set of genetic information in humans. It contains 3x10⁹ base pairs of DNA, dividing into nuclear DNA within 23 pairs of chromosomes and mitochondria DNA. Protein-coding DNA sequences constitute only ~1.1% of human genome and contain approximately 20000 protein-coding genes. The remaining human genome is the non-coding region that contains non-coding RNA, pseudo genes, introns, untranslated regions of mRNA, regulatory DNA sequences and repetitive DNA sequences. A typical gene contains promoter, exons, introns and 5’-/ 3’- untranslated regions. Genetic variations comprise both chromosome aberrations and DNA sequence changes. Variations can occur in coding or non-coding regions in human genome and they may be associated with diseases or abnormal phenotypes. There are several types of variation at DNA levels, e.g. single nucleotide polymorphisms (SNPs), variable
number tandem repeats (VNTR), copy number variations (CNV), insertions and deletions, all of which can be used as genetic markers to study the origin of diseases.

Linkage analysis and association study are two major approaches used in genetic studies of complex diseases. Linkage analysis is a family-based method to explore genomic regions transmitted along with the phenotype of interest. Classical linkage analysis is performed using microsatellites, which are highly polymorphic short tandem DNA sequence repeats. In population-based genetic association studies, candidate genes association study and genome-wide association study (GWAS) are most common-used approaches. SNP, in which a single nucleotide in DNA sequence is replaced by another, is the most common genetic variation in human genome and is routinely used as genetic-marker in association studies. The candidate genes approach is based on selection of SNPs within a gene, in which the gene production is involved in the pathogenesis of diseases or associated with phenotypes. Owning to the development of high-throughput SNP genotyping methodologies, genetic studies in complex diseases have moved into GWAS era nowadays. GWAS is a study focusing on common variants with minor allele frequency >5% and it allows association tests of more than 300,000 genetic markers (usually SNPs) across the genome with diseases or phenotypes.

1.2.3 Genetic studies in diabetes

In T1D, linkage analyses successfully identified human leukocyte antigen (HLA) region, which is located on human chromosome 6p21.3, to confer approximately 50% genetic susceptibility to T1D risk [31]. Moreover, more than 50 non-HLA genes are currently identified to be associated with T1D by GWAS and subsequent meta-analysis. Most of these genes are involved in immune response [32].

In T2D, the first two genetic associated variants located in the genes of peroxisome proliferator-activated receptor gamma (PPARG) and potassium inwardly-rectifying channel, subfamily J, member 11 (KCNJ11) were identified by the candidate gene approach [33]. PPARG encodes a receptor which is the target of an anti-diabetes drug thiazolidinedione. The common Pro12Ala polymorphism in PPARG is associated with reduced risk of T2D, indicating its protective effect in diabetes [34]. KCNJ11 encodes a membrane protein, which allows potassium influx into pancreatic beta cells. E23K variant in KCNJ11 is found to be associated with T2D by affecting insulin secretion [35]. KCNJ11 protein is associated with the sulfonylurea receptor, and stimulates insulin secretion [35].

Since the first report of GWAS in T2D in 2007, several GWAS in different populations and subsequent large-scale meta-analysis have discovered a number of genetic variants associated with T2D [36]–[44]. So far, more than 60 loci associated with T2D or glycemic traits have been identified.

Most of the identified loci are associated with impaired β cell function, e.g. the gene transcription factor 7-like 2 (TCF7L2). The TCF7L2 intronic variants has been recognized to be associated with T2D early in the pre-GWAS era by linkage study and confirmed later by several following GWAS [45]. TCF7L2 is believed to be the
strongest genetic risk factor for T2D and the combined odds ratio of T2D per copy of risk allele is 1.37 (95% CI= 1.31-1.43) [47]. TCF7L2 is involved in the Wnt signaling pathway that is important for β cell proliferation. Clinically, the TCF7L2 risk allele carriers had reduced insulin secretion [48]. Reducing TCF7L2 expression was found to decrease glucose induced insulin secretion in mouse islets and murine β cell lines [49]. To be mentioned, the variant of gene insulin-like growth factor 2 mRNA binding protein 2 (IGF2BP2) affects first-phase insulin secretion [48]. Reducing TCF7L2 expression was found to decrease glucose induced insulin secretion in mouse islets and murine β cell lines [49].

There are also some diabetes-associated loci related to impaired insulin sensitivity, e.g. fat mass and obesity-associated gene (FTO). FTO is the strongest obesity-associated gene and thus increases the risk of IR and T2D [51]. Of note, the variant rs35767, which is located 1.2 kb upstream of insulin-like growth factor 1 (IGF1), has been found to be associated with fasting insulin and HOMA-IR, manifesting the role of IGF1 in IR [52].

1.2.4 Genetic studies in diabetic nephropathy

Linkage studies in different populations have identified more than 17 chromosome regions related to proteinuria and/or ESRD in T1D and T2D [53]. The linkage of chromosome 3q with DN has been noticed in both T1D and T2D. Our group previously performed candidate gene genetic association studies of DN within chromosome 3q region, and identified several genetic polymorphisms in different genes associated with DN, including AdipoQ, Sox2, MCF2L2 and MME [54]–[57]. Interestingly, the T2D-related gene IGF2BP2 is also located in this chromosomal region.

Candidate gene association studies in DN are mostly focused on renin-angiotensin system, e.g. the genetic variants in the genes angiotensin 1-converting enzyme (ACE) and angiotensin II receptor type 1 (AGTR1) are reported to associate with DN in T1D [58], [59]. GWAS conducted in T1D-DN and T2D-DN have identified different nephropathy-associated genetic loci, indicating DN in T1D and T2D may not share the similar genetic background [60]–[62]. Nevertheless, the complexity of phenotypic definition of DN in T2D makes it more challenging to perform genetic study in T2D-DN and the only published GWAS in T2D-DN is conducted in African American population which has more risk to develop DN and ESRD, while GWAS in T1D-DN are most performed in Caucasian populations. Sandholm et al. recently revealed an association of one SNP in the AFF3 gene with ESRD in T1D patients, and further function study suggests that AFF3 influences renal tubular fibrosis via transforming growth factor-β pathway [60].

It is estimated that the identified genetic variants can explain around 80% heredity of T1D but only less than 15% in T2D [33]. The “missing heritability” can be explained by several factors. One possible reason is the limitation of GWAS. GWAS can only detect common variants with modest effect but not rare variants, thus deep sequencing using next-generation sequencing techniques in large population is necessary to search for less common or rare variants in diabetes and DN. Recently, using whole exome sequencing, Flannick et al. identified that 12 loss-of-function rare variants in solute
carrier family 30 (zinc transporter), member 8 (SLC30A8) gene are associated with reduced risk of T2D [63]. Other potential contributors of the missing heritability such as shared uterine or post-natal environment, gene-gene interaction, gene-environment interaction and latent epigenetic regulation are needed for further investigations [64].

1.3 EPICGENETIC STUDIES IN DIABETES AND DIABETIC NEPHROPATHY

1.3.1 The basics of epigenetics

The term epigenetics describes the phenomena of inherited gene expression alteration that occur independently of a change in DNA sequences [65]. Epigenetic changes are inheritable. The most-well known examples are X-chromosome inactivation and genomic imprinting, and both of them lead to monoallelic gene expression. X-chromosome inactivation is an epigenetic modification process to silence one of two copies of X chromosomes in females. Genomic imprinting terms the phenomena that only the non-imprinted allele in the imprinting gene is expressed, either inherited from the father (e.g. IGF-II) or mother (e.g. H19 or CDKN1C) [66].

It has been demonstrated that epigenome is dynamic changed in response to several exposures, such as aging, nutrition status, physical exercise, inflammatory, stress, etc. [67]–[72]. Epigenetic modulation is an important event not only in development, differentiation and tissue homeostasis, but also in the progression of diseases, especially in cancers, where the effects of epigenetic components have been best characterized. Cancer cells exhibit a global loss of DNA methylation and promoter hyper-methylation of tumor suppressor genes [73]. Epigenetic regulations are involved in the mechanisms of developmental reprogramming. Exposure to certain environmental events or nutrition status altered epigenetic modification since in utero or early in the childhood, and might have long term effects on adult disease in future [74], [75].

DNA methylation and histone modifications are two major types of epigenetic changes.

1.3.2 The basics of DNA methylation

DNA methylation occurs at the 5 position of cytosine residues, and dominantly in dinucleotide CpG sites in mammals (p stands for phosphate bond). DNA methylation regulates gene expression and transmits the specific expression pattern to daughter cells, such as in X-chromosome inactivation. There are two mechanisms of gene expression regulation by DNA methylation. The first mechanism is by affecting the binding of transcription factors in gene promoter. In human genome, there are regions called CpG islands, which are defined as in a genomic region with at least 200 base pairs, CG content over 50% and an observed-to-expected CpG ratio greater than 60% [76]. CpG islands are usually located in the 5’ ends of genes, including promoter, 5’-untranslated region and the first exon [77], thus the methylation at CpG islands will interfere the binding of gene transcription factors and repress the gene expression (Figure 2). The second mechanism was established with the identification of methyl-CpG binding protein (MBD), which binds to methylated CpGs and recruits protein complexes. These complexes contain histone-modifying enzymes and lead to gene silencing [78]. Two groups of methyltransferase are responsible for the activation of
DNA methylation. DNA-methyltransferase 1 (DNMT1) is required for methylation maintenance during DNA replication, while DNA-methyltransferase 3 alpha (DNMT3A) and DNA-methyltransferase 3 beta (DNMT3B) are necessary for de novo methylation [79], [80].

*Figure 2. The mechanism of gene expression regulation by DNA methylation*
1.3.3 The strategies of DNA methylation analysis

Similar to genetic studies, epigenetic studies can be performed in populations and within families. Unlike DNA sequencing, the tissue-specific and dynamic features of DNA methylation levels require more diverse study approaches in different biological materials. The advantages and disadvantages of current strategies for DNA methylation analysis are summarized in Table 2.

Table 2. Designs, approaches and biological materials used for DNA methylation analysis

<table>
<thead>
<tr>
<th>Design</th>
<th>Advantage</th>
<th>Disadvantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case-control study</td>
<td>Many cohorts exist</td>
<td>Difficult to control genetic and environmental confounders</td>
</tr>
<tr>
<td>Twins study</td>
<td>Control for genetics</td>
<td>Few large cohorts</td>
</tr>
<tr>
<td>Families study</td>
<td>Study potential inheritance</td>
<td>Few large cohorts</td>
</tr>
<tr>
<td>Longitudinal study</td>
<td>Determine causality</td>
<td>Time consuming</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Approach</th>
<th>Advantage</th>
<th>Disadvantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global DNA methylation analysis</td>
<td>General information of methylation in genome wide scale</td>
<td>Analysis of repeated sequence methylation Lack of gene specific information</td>
</tr>
<tr>
<td>Epigenome-wide association studies</td>
<td>Numerous CpG sites methylation information in genome wide scale</td>
<td>Higher cost Strict validation is needed</td>
</tr>
<tr>
<td>Specific gene DNA methylation analysis</td>
<td>Study candidates genes with potential biological functions</td>
<td>Less information on the studied genes</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Material</th>
<th>Advantage</th>
<th>Disadvantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell line</td>
<td>Intervention and mechanism study</td>
<td>In vitro experiment</td>
</tr>
<tr>
<td>Tissue</td>
<td>Gene specific methylation and expression can be analyzed</td>
<td>Difficult to access</td>
</tr>
<tr>
<td>Blood or saliva</td>
<td>Clinical accessible</td>
<td>Possible bias from mix cell types</td>
</tr>
</tbody>
</table>

1.3.4 DNA methylation studies in diabetes and diabetic nephropathy

Rakyan et al. performed epigenome-wide association study (EWAS) in purified CD 14+ monocyte from 15 T1D-discordant monozygotic twins and identified specific methylation variable positions associated with T1D susceptibility, which was subsequently replicated in another independent cohort [81]. In addition, DNA methylation in the proximal insulin gene promoter was reported to be associated with T1D [82].

In T2D, a monozygotic twin study showed a 10% increase of global methylation to be associated with an increase of 4.55 units of HOMA-IR [83]. Several recent studies were focused on characterizing the DNA methylation changes in specific tissues by candidate gene or EWAS approaches. In pancreatic islets, the DNA methylation
changes in the genes of INS, PDX1 and PPARC1A have been reported to contribute to the development of T2D [84]–[86]. By far, two EWAS have been performed in human pancreatic islets in 2012 and 2014, respectively [87], [88]. A number of CpG sites and genes were identified with differential DNA methylation in T2D islets. More than 100 genes with differential DNA methylation were also differentially expressed in T2D islets. Further functional analyses indicated identified genes affect insulin and glucagon secretion in β and α cells [88]. In addition, DNA methylation alteration is involved in the mechanism of exercise intervention in IR. Barees et al. reported that dynamic change of DNA methylation was able to active gene expression in human skeletal muscle in response to exercise [69].

Compared to specific tissues, measuring blood DNA methylation offers advantages of easy accessibility to samples and possibility of analysis in a large cohort. However, DNA methylation study from peripheral blood is more challenging owning to the heterogeneity of DNA source. Recently, a genome-wide methylation analysis demonstrated that increased methylation levels of gene HIF3A in both blood cells and adipose tissue were associated with increased BMI in a European population, suggesting that blood DNA methylation is able to reflect the phenotypic changes in specific tissues [89]. Genome-wide survey reveals that the low methylation levels of blood DNA at non-promoter genomic sites predispose to T2D [90]. Interestingly, a CpG site in the first intron of the FTO gene showed hypomethylation in T2D cases and the methylation alteration is under the control of FTO genetic variation, suggesting the potential interaction between epigenetic and genetic factors within disease-linked region [90], [91].

Nineteen prospective CpG sites were identified to be associated with DN risk in T1D patients, including one promoter CpG site in the UNCI3B gene. One intronic SNP in UNCI3B has recently been reported to be associated with DN, which resides within a linkage block of 23kb include the plausible promoter region, again indicating the potential interaction between genetic factors and epigenetic modulation [92].

1.4 IGF-IGFBP AXIS

1.4.1 Overview of IGF-IGFBP axis

The insulin-like growth factor (IGF)-IGF binding protein (IGFBP) axis consists of two insulin-like growth factors (IGF-I, IGF-II) and their binding proteins. Related to the axis, there are three IGF2 mRNA binding proteins (IGF2BP1, IGF2BP2 and IGF2BP3). The IGF-IGFBP axis plays important roles in growth and metabolism.

IGF-I is a peptide with 70 amino acid residues and shares structure homology with insulin [93]. Circulating endocrine IGF-I is predominately produced by the liver. A 75% reduction of serum IGF-I levels was observed in liver from IGF-I deficient mice [94]. Hepatic synthesis of IGF-I is mainly regulated by growth hormone (GH). GH deficiency patients have very low levels of GH, as well as low IGF-I due to the impaired hepatic synthesis. In GH resistance, GH receptors in the liver are unresponsive to GH and result in decreased IGF-I production, and subsequently lead to
hyper-secretion of GH [95]. Insulin and amino acids are needed for post receptor signaling to induce IGF-I synthesis [96]. IGF-I is also expressed in many other tissues, such as kidney, and acts as autocrine or paracrine factors. IGF-I exerts its effects of stimulating cell growth, survival and differentiation by binding to the IGF-I receptor. It also exhibits its insulin-like effects in metabolism by binding to insulin receptor or the hybrid receptor of IGF/insulin receptors, although IGF-I binds to insulin receptor with only one-hundredth affinity compared to insulin [97]. IGF-I levels peak at puberty and then decline by age [98]. Unlike insulin, which is largely unbound to any molecule, only around 1% of total IGF-I in circulation is in an “unbounded” form.

IGF-II shares about 70% amino acid identity with IGF-I. The igf2^-/- mice are born 40% smaller than control mice, but the growth rate after birth is similar between knockout and control animals, indicating IGF-II is more important for fetal growth [99],[100]. IGF-II can bind to IGF-I receptor, mannose-6-phosphate (M6P)/IGF-II receptor and hybrid receptor. The M6P/IGF-II receptor mediates the action of IGF-II on exocytosis in insulin secreting cells. IGF-II can stimulate the insulin release from β-cells at basal concentration of glucose, but inhibit glucose-induced insulin release [101],[102].

**IGF binding proteins**

Most of the IGFs (approximately 99%) in circulation bind to a group of binding proteins, termed as IGF binding proteins (IGFBPs). There are six IGFBPs (IGFBP-1 to -6) with conserved structure binding to IGFs with high affinity, while IGFBP-7 binds to IGFs with low affinity [103]. IGFBPs are present in body fluids and tissues, varying in molecular size, biological function and hormone regulation. IGFBPs are involved in several biological functions, e.g. prolong the IGFs half-life, store and transport IGFs, modulate the activity of IGFs, as well as IGF-independent actions on cellular proliferation and migration [104]. The IGFBPs generally decrease the bioactivity of IGFs by competing with IGF receptors for IGF binding. Posttranslational modifications of IGFBPs can affect their IGF binding affinity and thus regulate IGFs actions. For example, phosphorylation of IGFBP-1 increases its IGF binding affinity and promotes the inhibitory effect of IGFBP-1 on IGF actions [105]. Serum proteolysis of IGFBPs, particularly proteolysis of IGFBP-3 can reduce the IGF binding affinity and thereby increase the IGF bioactivity [106].

**IGFBP-3** is the most abundant IGFBP in circulation. It binds to IGFs together with an acid-labile subunit (ALS) to form a ternary complex. This 150 kDa complex is not able to cross the vascular endothelium. Both IGFBP-3 and ALS are mainly produced from liver and regulated by GH [107]. An in vitro study suggested that IGFBP-3 had an IGF-independent function that inhibited cellular proliferation in breast cancer cells [108].

**IGFBP-1** is thought to be most metabolically regulated IGFBP. Hepatic IGFBP-1 production, which is the major source of circulating IGFBP-1, is inhibited by insulin and stimulated by glucagon, cortisol, fasting or cytokines [109], [110]. In the circulation, similar to IGFBP-2, -4, and -6, IGFBP-1 forms a binary complex with IGFs, which is able to cross the vessel wall to transport IGFs into target tissues. IGFBP-1 has been considered to be the acute regulator of IGF-I bioavailability by regulating the “unbounded” free IGF-I levels in vivo [111]. It has demonstrated that the
Arg-Gly-Asp (RGD) conserved sequence in IGFBP-1 binds to the alpha 5 beta 1 integrin (α5β1 integrin) and stimulates the cell migration and proliferation, indicating IGFBP-1 has IGF-independent effects [112].

**IGFBP-7** was identified as an additional IGFBP family member in 1996. Oh *et al.* showed that IGFBP-7 contains an IGFBP conserved motif at the NH2 terminal and is expressed in different tissues [103]. IGFBP-7 can bind to IGF-I, IGF-II and insulin. Compared with the other six IGF binding proteins, IGFBP-7 binds the IGFs with 5- to 25-fold lower affinity, whereas it binds insulin with 500-fold higher affinity [113].

Three IGF-II mRNA binding proteins IGF2BP1, IGF2BP2 and IGF2BP3 belong to a conserved family of RNA-binding protein, and they are important for RNA localization, translation and stability [114]. The expression of IGF2BPs is most recognized in embryo development. However, postnatal expression of IGF2BPs has also been observed. IGF2BP2 mRNA expression in several adult tissues has been detected, including brain, gut, kidney, liver, lung, muscle and pancreases in human or mouse [115].

**Figure 3. IGF-I, IGFBP-1 and IGFBP-3 in liver, circulation and target tissues**
1.4.2 IGF-IGFBP axis in glucose homeostasis and diabetes

The maintenance of glucose homeostasis depends on the balance between glucose production and utilization, and requires the complex interplays between several hormones. In addition to insulin, the IGF-IGFBP axis plays an important role in glucose homeostasis.

IGF-I stimulates glucose uptake in skeletal muscle directly via IGF-I receptor, although its potency is only 4-7% of insulin [116]. Plasma IGF-I levels are independently correlated to insulin sensitivity, which has been suggested to be a marker for IR [117]. Treatment with recombinant IGF-I has been shown to reduce blood glucose in healthy individuals [118], as well as in T1D and T2D patients [119], [120]. Furthermore, IGF-I can suppress insulin secretion in healthy adults in short term [121], and improve insulin sensitivity in both short and prolong terms [122]. As a mitogen, IGF-I also stimulates the proliferation of pancreatic β cells [123].

IGFBP-1 acutely regulates blood glucose by binding to free IGF-I [111]. Insulin inhibits the hepatic production of IGFBP-1[124]. Circulating IGFBP-1 is inversely correlated with insulin under normal condition. Furthermore, fasting serum IGFBP-1 is inversely correlated to hepatic insulin sensitivity using measured using the euglycemic hyperinsulineamic clamp. Therefore, IGFBP-1 is a reliable marker of insulin secretion and insulin sensitivity [98] [125].

IGFBP-2 is secreted from liver, as well as white pre-adipocytes during adipogenesis [126]. Hedbacker et al. demonstrated that IGFBP-2 treatment can improve hepatic insulin sensitivity in diabetic ob/ob mice [127], suggesting the important role of IGFBP-2 in glucose metabolism regulation. IGFBP-3 is the most abundant circulating IGFBP and the metabolic actions of IGFBP-3 are mostly opposite to IGF-I [128]. As an insulin antagonist, IGFBP-3 decreases the peripheral glucose uptake [128].

The IGF-IGFBP axis is disturbed in diabetes. Newly diagnosed T1D patients have high levels of GH but low levels of IGF-I, due to the low portal insulin levels and impaired GH receptor function, suggesting GH resistance and relatively IGF-I deficiency in T1D [116][124]. Meanwhile, IGFBP-1 is significantly elevated in T1D because of its increased hepatic production [124]. In response to insulin therapy, IGF-I increases whereas IGFBP-1 decreases [129].

In T2D, the IGFBP-1 levels vary due to the changes of insulin levels at different stages of disease. Low levels of IGFBP-1 predict the development of impaired glucose regulation [130]–[132]. High levels of IGFBP-1 have been shown to be associated with increased risk of cardiovascular mortality and morbidity in T2D patients with acute myocardial infarction [133]. The evidence implies IGFBP-1 as a predictor for the risk of diabetes and cardiovascular diseases. It has been demonstrated that high circulating IGF-I levels were associated with reduced risk of development of impaired glucose tolerance and T2D among the subjects with normal glucose concentrations at baseline [134]. In individuals with low IGFBP-1 levels, IGF-I concentration was inversely correlated with 2h-glucose concentration, indicating the interaction between IGF-I and...
IGFBP-1 is important in glucose regulation [134]. Clinical observation has demonstrated that serum IGFBP-2 levels were reduced in obese subjects and T2D patients when compared with lean subjects[135]. Low levels of IGFBP-2 have been found to be associated with an increased risk of metabolic syndrome in T2D patients [136].

1.4.3 **IGF-IGFBP axis in diabetic nephropathy**

IGF-IGFBP axis also plays a role in the development of DN. In diabetic animal models, the local accumulation of IGF-I peptide in kidney has been observed before renal hypertrophy, which is the early stage of DN [137], [138]. Furthermore, the decreased *IGF1* mRNA expression levels detected in the kidney indicated that the increased renal IGF-I protein was probably not due to local IGF-I production. It is suggested that the local interaction of IGF-I and IGFBP-1 alters the IGF-I activity and leads to renal hypertrophy [139], [140]. Clinical observation has shown that T1D patients with microalbuminuria have decreased IGF-I and increased IGFBP-1 serum levels compared with patients with normalalbuminuria [141]. It was demonstrated that in T2D patients, high IGFBP-2 concentration at baseline was associated with decreased renal function over an 8-year period, indicating that IGFBP-2 could be used to predict the development of DN in T2D [142].
1.4.4 Genetic studies in IGF-IGFBP axis

General information of genes in the IGF-IGFBP axis is shown in Table 3. Previous studies have identified that several polymorphisms in the genes of IGF-IGFBP axis are associated with diseases and phenotypes (summarized in Table 4).

Table 3. The general information of the genes in IGF-IGFBP axis

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Gene size (bp)</th>
<th>Chromosomal location</th>
<th>Biological function</th>
</tr>
</thead>
<tbody>
<tr>
<td>IGF1</td>
<td>84,779</td>
<td>12q23.2</td>
<td>It is similar to insulin in function and structure and mediates growth and development.</td>
</tr>
<tr>
<td>IGF2</td>
<td>20,492</td>
<td>11p15.5</td>
<td>It plays a role in fetal development. It is an imprinted gene, expressed only from the paternal allele.</td>
</tr>
<tr>
<td>IGFBP1</td>
<td>5,312</td>
<td>7p12.3</td>
<td>It is considered to be the acute regulator of IGF-I bioavailability and a marker of peripheral insulin sensitivity. It stimulates cell migration and proliferation by its IGF independent effect.</td>
</tr>
<tr>
<td>IGFBP2</td>
<td>31,609</td>
<td>2q35</td>
<td>It is the principal IGFBP secreted by white adipose tissue. It is regulated by leptin and has anti-diabetic effect.</td>
</tr>
<tr>
<td>IGFBP3</td>
<td>9,630</td>
<td>7p12.3</td>
<td>It is the most abundant IGFBP in circulation, binds IGF together with ALS and forms a ternary complex to store IGF.</td>
</tr>
<tr>
<td>IGFBP4</td>
<td>14,308</td>
<td>17q21.2</td>
<td>Limited information. IGFBP-4 and -6 binds to IGFs as a binary complex, whereas IGFBP-5 binds to IGFs as ternary complex together with ALS.</td>
</tr>
<tr>
<td>IGFBP5</td>
<td>23,445</td>
<td>2q35</td>
<td></td>
</tr>
<tr>
<td>IGFBP6</td>
<td>4,910</td>
<td>12q13.13</td>
<td></td>
</tr>
<tr>
<td>IGFBP7</td>
<td>79,613</td>
<td>4q12</td>
<td>It binds IGFs with low affinity, and binds insulin with higher affinity than other IGFBPs.</td>
</tr>
<tr>
<td>IGF2BP1</td>
<td>58,734</td>
<td>17q21.32</td>
<td>IGF2BPs are mainly expressed in the embryo. IGF2BPs bind specific target mRNAs, including IGF2 mRNA, control the mRNA localization, stability and translation.</td>
</tr>
<tr>
<td>IGF2BP2</td>
<td>181,318</td>
<td>3q27.2</td>
<td></td>
</tr>
<tr>
<td>IGF2BP3</td>
<td>160,259</td>
<td>7p15.3</td>
<td></td>
</tr>
</tbody>
</table>
Table 4. The genetic polymorphisms in the IGF-IGFBP axis and their associations

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Genetic polymorphism</th>
<th>Association</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>IGF1</td>
<td>rs35767 (C/T)</td>
<td>Associated with fasting insulin and HOMA-IR index.</td>
<td>Josee et al. 2010 [52]</td>
</tr>
<tr>
<td></td>
<td>rs5742692 (G/A)</td>
<td>Associated with fasting, 2-h insulin levels as well as insulin sensitivity.</td>
<td>Hu et al. 2010 [143]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Associated with circulating IGF-I and insulin sensitivity.</td>
<td>Gaia et al. 2013 [144]</td>
</tr>
<tr>
<td>IGFBP1</td>
<td>rs1065780 (G/A)</td>
<td>Associated with adult height.</td>
<td>Okada et al. 2010 [145]</td>
</tr>
<tr>
<td></td>
<td>rs3828998 (T/C)</td>
<td>Associated with a reduced prevalence of DN in T2D.</td>
<td>Stephens et al. 2005 [146]</td>
</tr>
<tr>
<td></td>
<td>rs3793344 (A/G)</td>
<td>Associated with IGF1 activity and lipid levels in adolescents.</td>
<td>Mong et al. 2009 [147]</td>
</tr>
<tr>
<td>IGFBP3</td>
<td>rs2854744 (A/C)</td>
<td>Associated with a reduced risk of pancreatic cancer.</td>
<td>Suzuki et al. 2008 [149]</td>
</tr>
<tr>
<td></td>
<td>rs13223993 (A/G)</td>
<td>Associated with increased maternal glucose concentrations in the third trimester of pregnancy and placental IGF-II contents at birth.</td>
<td>Petry et al. 2011 [150]</td>
</tr>
<tr>
<td></td>
<td>rs680 (C/G/T)</td>
<td>Associated with polycystic ovary syndrome.</td>
<td>San Millan et al. 2004 [151]</td>
</tr>
<tr>
<td></td>
<td>rs6578987 (C/T)</td>
<td>Associated with IGF-II concentration and longitudinal weight changes.</td>
<td>Narayanan et al. 2013 [152]</td>
</tr>
<tr>
<td></td>
<td>rs7924316 (G/T)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>rs10770125 (A/G)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>rs680 (C/G/T)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>rs6578987 (C/T)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>rs7924316 (G/T)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>rs10770125 (A/G)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGFBP5</td>
<td>rs9341234 (C/T)</td>
<td>Associated with IGFBP2 mRNA and protein levels in visceral adipose tissue and insulin resistance in T2D.</td>
<td>Chistiakov et al. 2012 [156]</td>
</tr>
<tr>
<td></td>
<td>rs3276 (A/G)</td>
<td>Associated with body fat and T2D-relative quantitative traits.</td>
<td>Li et al. 2009 [157]</td>
</tr>
<tr>
<td></td>
<td>rs11575134 (A/G)</td>
<td>Associated with T2D.</td>
<td>Saxena et al. 2007 [37]</td>
</tr>
<tr>
<td></td>
<td>rs4402960 (G/T)</td>
<td></td>
<td>Scott L et al. 2007 [38]</td>
</tr>
<tr>
<td></td>
<td>rs3847627 (A/G)</td>
<td></td>
<td>Zeggini et al. 2007 [39]</td>
</tr>
<tr>
<td></td>
<td>rs11705701 (A/G)</td>
<td></td>
<td>Sanghera et al. 2008 [153]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Takeuchi et al. 2009 [154]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Han et al. 2010 [155]</td>
</tr>
</tbody>
</table>
2 AIMS

2.1 GENERAL HYPOTHESIS

Diabetes and DN are complex diseases reflecting a complex interplay between genetic and non-genetic factors. The IGF-IGFBP axis plays an important role in the development of diabetes and DN, and recent reports have demonstrated that genetic polymorphisms in the genes of this axis are associated with diabetes and DN. However, the information of epigenetic studies is very limited. We hypothesized that genetic variations and epigenetic alterations in our candidate genes in the IGF-IGFBP axis are associated with diabetes and DN.

2.2 AIMS

In this thesis, we selected four genes from the IGF-IGFBP axis including \textit{IGF1}, \textit{IGF2}, \textit{IGFBP1} and \textit{IGF2BP2} to evaluate their genetic and epigenetic associations with diabetes and DN. In parallel, we analyzed the serum protein levels.

Study I: To investigate the DNA methylation alteration and genetic variation of the \textit{IGF1} gene in relation to IGF-I serum levels in Swedish males with T2D, compared to controls.

Study II & III: To analyze \textit{IGFBP1} DNA methylation levels in relation to IGFBP-1 serum levels in T2D, T1D with or without DN and controls in Swedish males.

Study IV: To evaluate the genetic effects of \textit{IGF2BP2} and \textit{IGF2} in diabetes and DN.
3 MATERIALS AND METHODS

3.1 SUBJECTS

3.1.1 Stockholm diabetes prevention program (SDPP)

The SDPP is a prospective population based study. The population consists of residents from five municipalities in the Stockholm County: Sigtuna, Tyresö, Upplands-Bro, Upplands Väsby and Värmdö. The subjects comprise both males and females, and my thesis work focused on the male subjects only. At baseline, 3128 male participants aged 35-55 years who were without known diabetes were enrolled into the study in 1992-1994. After 8-10 years, the baseline study group, except those with newly diagnosed diabetes at baseline, was invited for a follow-up study. In total, 2383 male participants from baseline group were included into the follow-up study. All of them were either normal glucose tolerance (NGT) or pre-diabetes at baseline. The individuals underwent a standard OGTT, body measurements and answered a questionnaire about life style factors at both baseline and follow-up.

In this thesis, all DNA samples selected for methylation studies are from SDPP follow-up study, including 164 cases (T2D) and 242 controls (NGT). Among the cases, 75 individuals were diagnosed diabetes during the time between baseline and follow-up studies, and they were treated with advices on physical exercise and diet control (29.3%), oral anti-diabetic drugs (OAD) (53.3%), insulin (5.3%) or a combination of these alternatives (4.0%). The other 89 cases were diagnosed in the follow-up study. Information of smoking status, physical activity levels and alcohol consumption in all participants were recorded based upon questionnaires. The information regarding family history of diabetes (FHD) was collected from all subjects. FHD is defined as having at least one first-degree relative (parents or siblings) or at least two second-degree relatives (father’s or mother’s parents or siblings).

3.1.2 Genetics of kidneys in diabetes study (GoKinD)

The GoKinD collection was supported by the Juvenile Diabetes Research Foundation in collaboration with the Joslin Diabetes Center, George Washington University, and the United States Centers for Diabetes Control and Prevention.

Among the GoKinD population, the majority of subjects were of European descents (n=1139), and approximately 8.5% of subjects were Native Indians, African-, Hispanic- and Asian Americans. To avoid confounding results caused by the ethnic groups, this small proportion of subjects was excluded for genetic study. All of them were the patients with T1D diagnosed before 31 years of age and treated with insulin within one year of diagnosis. T1D subjects with DN (cases, n=559, males 304/ females 255) had either persistent proteinuria, defined by a urinary albumin/creatinine ratio (ACR) $\geq$ 300 μg/mg in two of the last three measurements taken at least 1 month apart, or ESRD (dialysis or renal transplant). Of the cases, 70.6% (n=408) were ESRD. T1D subjects without DN (controls, n=580, males 231/ females 349) had T1D for at least 15 years and normal albuminuria, defined by an albumin to creatinine ratio < 20 μg/mg in two of the last three measurements taken at least 1 month apart, without ever having been treated with ACE inhibitors or angiotensin receptor blockers.
3.1.3 Subjects from Czech Republic

This cohort was collected in the Third Department of Internal Medicine, Charles University and General Faculty Hospital in Prague, Czech Republic (2002–2010), including 1399 subjects of European descents in total. Among them, 339 (males 106 /females 233) were non-diabetic control subjects, 243 (135/108) were T1D patients and 817 (420/397) were T2D patients. T1D was diagnosed before 35 years of age and the time to definitive insulin therapy was ≤1 year. The patients with T2D were diagnosed without age limitation and treated with OADs, insulin or both. All patients with diabetes were divided into subgroups according to diabetes type and renal function (T1D without DN, T1D with microalbuminuria, T2D without DN, T2D with microalbuminuria, and T2D with DN). Absence of DN was assumed when the subjects had persistent normal albuminuria (albuminuria <30 mg/24 h or <20 µg/min or <20 mg/l or ACR <2.5 mg/mmol). Presence of microalbuminuria was defined by urinary albumin excretion rate (AER) 30–300 mg/24 h or 20–200 µg/min or 20–200 mg/l or ACR 2.5–25 mg/mmol. Presence of DN was defined either by persistent proteinuria (>300 mg/24 h or >200 µg/min or >200 µg/min or ACR>25 mg/mmol) or chronic kidney disease (glomerular filtration rate GFR <60 ml/min) or ESRD.

3.1.4 Swedish subjects with type 1 diabetes and diabetic nephropathy

A total of 536 Swedish T1D patients were collected in the Department of Endocrinology, Karolinska University Hospital. Urinary AER 20-200 µg/min in at least two consecutive overnight samples was considered as micro-albuminuria, while AER >200 µg/min in at least two consecutive overnight samples as macro-albuminuria. 51 (25 males/26 females) T1D patients had macro-albuminuria were classified as the cases of T1D with DN, including two T1D patients who received renal replacement therapy, while 296 (160 males /136 females) T1D patients with persistent normal albuminuria were grouped as the controls of T1D without DN. In addition, 189 (119 males/70 females) T1D patients with normal albuminuria or historic micro-albuminuria had medical treatments with angiotensin-converting-enzyme inhibitor (ACEI)/angiotensin II receptor blockers (ARB).

3.2 ANIMAL MODEL

3.2.1 db/db mice

The db/db mouse is a well-characterized and intensely investigated animal model for DN study. It was identified as an obese model in Jackson Labs in 1996. This animal model has impaired leptin signaling resulting from a point mutation in the gene of leptin receptor, and exhibits persistent hyperphagia, obesity, high levels of insulin, leptin and blood glucose.

In db/db mice on C57BLKS/J background, hyperinsulinemia can be noted by 10 days of age and their blood glucose levels are slightly increased at 1 month [158]. As an animal model for DN, db/db mice share the similar phenotypes with humans in kidney hypertrophy, glomerular enlargement, mesangial matrix expansion and albuminuria. Increased glomerular size starts to occur during the early stages of diabetes (around 8 weeks) in db/db mice. After 16 weeks, increased mesangial matrix expansion in db/db
mice can be consistently observed [159–161]. Compared to age-matched heterozygous littermate, the albumin excretion rates in db/db mice are 8- to 62-fold higher, but unlike in humans, the degree of albuminuria doesn’t consistently increase with diabetes duration [158].

Given the above features of db/db mice, in study IV, we extracted mRNA and protein from kidney tissues in db/db mice and lean lC57BLKS/J control mice, at the ages of 5 weeks and 26 weeks, to detect igf2bp2 gene expression at mRNA and proteins levels.

3.3 METHODS

3.3.1 Taqman allelic discrimination

Taqman allelic discrimination is one of the standard methods to detect variants of a single nucleic acid sequence. Two probes with the incorporation of minor groove binder (MGB) at the 3’ end are used to target two alleles. Each probe has a unique reporter dye at 5’ end to distinguish two alleles, and also a quencher dye attached at 3’ end. During the extension by primer, exonuclease cleavage of an allele-specific 5’ reporter dye generates increased fluorescence intensity and then determined by laser detection. Our genotyping experiments were carried out with TaqMan Allelic Discrimination protocol by using ABI 7300 system (ABI, Foster City, CA, USA). For quality control, DNA samples are distributed randomly across plates with the cases and controls per PCR plate. Negative controls (Universal-mixture blanks) are included on each plate. A subset of randomly selected samples representing ~20% of the study subjects is replicated.

3.3.2 Bisulfite pyrosequencing DNA methylation analysis

Bisulfite sequencing is widely used as the “golden standard” method for DNA methylation measurement [162]. Incubation of the DNA samples with sodium bisulfite allow un-methylated cytosine convert to uracil, leaving the methylated cytosine unchanged (Figure 4A). After amplifying target DNA sequence by PCR, the ration C (methylated cytosine) to T (un-methylated cytosine) can be used to evaluate the methylation status at a given position, and to transfer epigenetic information to sequence information for qualification by pyrosequencing.

Pyrosequencing is a sequencing-by-synthesis method to detect the light signal produced by nucleotide incorporation and subsequent pyrophosphate release. Therefore, pyrosequencing-based DNA methylation analysis is a sensitive and quantitative method. It allows accurate measurements of several adjacent CpG sites in one simple reaction [162]. Practical procedures of bisulfite pyrosequencing DNA methylation measurement are divided into several steps as described in Figure 4B.

First, genomic DNA (usually 500ng per time) was treated by sodium bisulfite and then cleaned up using EpiTect Bisulfite Kit (Qiagen), which gives complete conversion of un-methylated cytosine to uracil and purification.

Second, PCR primers were designed according to the converted DNA sequence and one of the primers (either forward or reverse primer) is biotin labeled. The size of the
amplification product is up to 350 bp. Pyrosequencing primer was also designed to analyze the target sequence with CpG sites of interest.

Third, DNA template for sequence was prepared by PCR amplification using the PyroMark PCR Kit (Qiagen). A minimum of 10 ng DNA was necessary to obtain enough signals for methylation detection. 45 to 50 cycles were required to exhaust the free biotinylated primer and obtain sufficient PCR product.

Finally, analysis was performed with the platform of PyroMark Q96 ID pyrosequencing system (Biotage, Uppsal, Sweden). Biotin labeled DNA was captured by streptavidin-coated sepharose beads for purification and then released into annealing buffer together with sequencing primer for sequencing reaction. The methylation levels at CpG sites were calculated by PyroQ-CpG software (Biotage). Completely methylated or completely un-methylated bisulfite converted DNAs, and untreated, unmethylated genomic DNA standards (Biotage) were used for standardization and reliable control reactions for methylation analysis.

Figure 4A. Bisulfite conversion of cytosine to uracil

Figure 4B. The procedures of Bisulfite pyrosequencing DNA methylation
3.3.3 Serum IGF-I and IGFBP-1 measurements

Serum levels of IGF-I were determined by radio immunoassay (RIA) after separating from IGFBPs by acid ethanol extraction and cryo-precipitation. Truncated IGF-I that lacks the N-terminal tripeptide glycine-proline-glutamate and has low affinity for IGFBPs was used as a radioligand, in order to minimize the interference from the remaining IGFBPs. The intra- and inter- assay coefficient of variations were 4% and 11%, respectively [163]. IGFBP-1 concentration in serum was determined by an in-house RIA using a polyclonal antibody as previously described. IGFBP-1 isolated from human amniotic fluid was used as standard. The intra- and inter- assay coefficient of variations are 3% and 10%, respectively [164].

3.3.4 Insulin and HbA1c measurements

Immuno-reactive insulin was assayed by an in house RIA, using a polyclonal antibody and human insulin as a standard. HbA1c was determined using immunologic MonoS assay, Unimate (Roche Diagnostics, Basel, Switzerland). To convert HbA1c MonoS to NGSP, the formula NGSP= 0.92*MonoS+1.33 was used.

3.3.5 Analyses of mRNA expression and protein levels

Kidney tissues were isolated from db/db and control mice and submerged in RNA later solution (Sigma-Aldrich, Buche, Germany). Total RNA was extracted from kidney tissues using RNAeasy Mini kit according to manufacturer’s instructions (Qiagen, Hilden, Germany). cDNA reverse transcription was performed using QuantiTect Reverse Transcription kit (Qiagen). Taqman real-time PCR was performed with ABI 7300 real-time PCR system to measure mRNA expression. Primers and assays specific for studied genes were purchased from Applied Biosystems (USA).

We extracted protein from mice’ kidney tissues using lysis RIPA buffer with protease inhibitor cocktail. The protein concentration was qualified using a protein assay (Bio-Rad laboratories, California, USA), electrophoresed with SDS 7.5% PAGE, transferred to nitrocellulose membrane and blocked with 5% non-fat milk. Primary antibody (Novus Biologicals, Cambridge, UK) was added in 1:500 dilutions and incubated overnight at 4 °C. After washing with PBS buffer, the secondary antibody (Novus) was added at 1:3000 concentration and incubated for one hour in room temperature. Bound antibody was detected by ECL western blotting system (GE Healthcare, USA).

3.4 STATISTICS

All data were analyzed using PASW statistic program (SPSS 20.0, Chicago, Illinois, USA). P-values less than 0.05 were considered as significant. Hardy-Weinberg equilibrium was tested in controls. Genotypes and allele frequencies of SNPs were analyzed between cases and controls by chi-square test. Continues variables between groups were compared by unpaired t-test or one-way ANOVA followed with Turkey post hoc test. Covariate-adjusted generalized models were used when adjusting for co-variables. Linear regression analyses were used to examine the relation between two continuous variables. Non-normally distributed continuous variables were log-transformed before analysis to improve the normal distribution. Gene-gene interaction
was detected using generalized multifactor dimensionality reduction (GMDR) program. GMDR is a nonparametric and genetic model-free alternative to linear or logistic regression program for detection of gene-gene or gene-environmental interaction.
4 RESULTS

4.1 STUDY I

IGF1 DNA methylation and serum IGF-I levels

The methylation levels of three CpG sites nearby rs357676 in the promoter of the IGF1 gene were evaluated. Compared with NGT subjects, T2D patients had significantly increased DNA methylation levels at CpG site P3 (84.8% vs. 74.2%, P<0.001). No difference was observed between NGT and T2D in DNA methylation levels at CpG sites P1 and P2. Furthermore, newly diagnosed and treated T2D patients had the similar methylation levels at three CpG sites. Circulating levels of IGF-I were found to be lower in T2D patients compared with NGT subjects (152 µg/l vs. 165 µg/l, P=0.029 adjusted for age), and the mean IGF-I SD score was lower in T2D patients than in NGT subjects (0.01 vs. 0.28, P=0.011).

Epigenotypic analyses according to SNP rs35767 in the IGF1 gene

The allele T in rs35767 removes a CpG site in the IGF1 promoter region. We found that the carriers with CC, CT and TT genotypes of rs35767 among NGT subjects had decreasing DNA methylation levels from 93.1%, 47.6% to 1.5%, and similar distribution of DNA methylation levels in the carriers with CC, CT and TT genotypes was seen among T2D patients (93.2%, 46.4% and 1.3%, P<0.001). There was no difference of the methylation levels in any of the genotyping groups between NGT and T2D. Meanwhile, there was an increased tendency of serum IGF-I levels in T allele carriers in NGT subjects (CC: 161 µg/l vs. CT: 168 µg/l vs. TT: 188 µg/l, P=0.259).

4.2 STUDY II & III

IGFBP1 DNA methylation and serum protein variations between NGT and diabetes

We analyzed DNA methylation levels at six CpG sites (referred as P1-P6) of the IGFBP1 gene. Compared with those in NGT subjects, the IGFBP1 DNA methylation levels were significantly decreased at five of the six CpG sites in T1D patients, but increased at all six CpG sites in T2D patients. Combining all six CpG sites together, the men levels of IGFBP1 DNA methylation were significantly lower in T1D patients, but higher in T2D subjects than those in NGT individuals. Meanwhile, serum levels of IGFBP-1 in T1D patients were increased, but decreased in T2D subjects. Data of the IGFBP1 DNA methylation and serum IGFBP-1 levels in Swedish men with NGT, T1D and T2D are summarized in Table 5.
Table 5. A summary of IGFBP1 DNA methylation and serum IGFBP-1 levels in Swedish men with normal glucose tolerance, type 1 and type 2 diabetes

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>IGFBP1 DNA methylation levels (%)</th>
<th>P-value*</th>
<th>IGFBP-1 serum levels (µg/l)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>T2D</td>
<td>164</td>
<td>20.0 (19.5-20.5)</td>
<td>P&lt;0.001</td>
<td>18 (16-20)</td>
<td>0.014</td>
</tr>
<tr>
<td>NGT</td>
<td>242</td>
<td>16.9 (16.4-17.1)</td>
<td>-</td>
<td>24 (21-26)</td>
<td>-</td>
</tr>
<tr>
<td>T1D</td>
<td>304</td>
<td>15.6 (15.1-16.1)</td>
<td>P&lt;0.001</td>
<td>31 (28-35)</td>
<td>0.007</td>
</tr>
</tbody>
</table>

Data were presented as means (95%CI). NGT: normal glucose tolerance; T1D: type 1 diabetes; T2D: type 2 diabetes; P-values were from comparison tests of NGT vs. T1D or T2D and * adjusted for age.

**IGFBP1 DNA methylation changes and serum protein variation between T1D with and without DN**

We further analyzed DNA methylation levels of *IGFBP1* gene and serum IGFBP-1 levels in DN. There was no difference in *IGFBP1* DNA methylation levels between T1D with and without DN in neither males (16.4% vs. 15.4%, P=0.186) nor females (15.9% vs. 15.6%, P=0.604). Compared with T1D patients without DN, however, the IGFBP-1 serum levels in T1D patients with DN were significantly increased in both males (52 µg/l vs. 28 µg/l, P=0.021) and females (71 µg/l vs. 33 µg/l, P=0.003). However, no difference in serum IGFBP-1 levels was found between males and females in T1D with neither DN (52 µg/l vs. 71 µg/l, P=0.17) nor the patients without DN (28 µg/l vs. 33 µg/l, P=0.28).

**IGFBP1 DNA methylation and serum protein levels according to BMI**

The *IGFBP1* DNA methylation levels were not dependent of BMI in T1D, T2D and NGT subjects, but the serum levels of IGFBP-1 were associated with BMI. The lean individuals (BMI<25 kg/m²) in each group of NGT, newly diagnosed and treated T2D patients had significantly higher serum IGFBP-1 compared with the subjects with overweight/obesity (BMI≥25 kg/m²) (29 vs. 22 µg/l in NGT, P=0.022; 29 vs. 16 µg/l in newly diagnosed T2D, P=0.011; 28 vs. 16 µg/l in treated T2D, P=0.020). Similarly, all T1D patients were divided into four groups according to BMI, i.e. low weight group (BMI <18.5 kg/m²), normal weight group (18.5-25.0 kg/m²), overweight group (25.0-30.0 kg/m²) and obese group (≥30.0 kg/m²). Circulating IGFBP-1 levels were decreased gradually according to BMI (51, 39, 30 and 28 µg/l, P=0.012), whereas the *IGFBP1* DNA methylation levels were similar among the four groups (16.0%, 15.9%, 15.6% and 16.0%, P=0.867).

**IGFBP1 DNA methylation and serum IGFBP-1 levels according to family history of diabetes**

NGT subjects with and without FHD had similar *IGFBP1* DNA methylation levels (16.9% vs. 17.0%, p=0.895) but the NGT subject with FHD had significantly lower serum protein levels (19 vs. 25 µg/L, P=0.018). In newly diagnosed T2D patients, compared with those without FHD, the patients with FHD had significantly increased *IGFBP1* DNA methylation levels (20.3% vs. 18.6%, P=0.014).
4.3 STUDY IV

The association of IGF2BP2 rs4402960 with T2D and diabetic nephropathy

The association of IGF2BP2 rs4402960 with T2D in the Czech population was replicated in this study. The frequency of T allele in the patients with T2D was higher compared with control subjects without diabetes (0.34 vs. 0.29, P=0.025, OR=1.26, CI 95% 1.03-1.54). Furthermore, there was a difference of T allele frequencies between T1D patients with and without DN among male subjects of the GoKinD population (0.30 vs. 0.35, P=0.047) but not in females (0.35 vs. 0.34, P=0.796). The association of IGF2BP2 rs4402960 with DN in male patients with T1D was still remained when we combined GoKinD, Czech and Swedish populations (P=0.030, OR=0.73, 95%CI 0.54-0.97) (Table 6).

Table 6. Association of IGF2BP2 rs4402960 with diabetic nephropathy in type 1 diabetes

<table>
<thead>
<tr>
<th>Original</th>
<th>Gender</th>
<th>Sample size (cases/controls)</th>
<th>MAF (cases/controls)</th>
<th>Significance P-value; OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GoKinD</td>
<td>Male</td>
<td>535 (304/231)</td>
<td>T 0.30/0.35</td>
<td>0.037; 0.69 (0.49-0.98)</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>604 (255/349)</td>
<td>T 0.35/0.34</td>
<td>0.845</td>
</tr>
<tr>
<td>GoKinD</td>
<td>Male</td>
<td>801 (398/403)</td>
<td>T 0.30/0.34</td>
<td>0.030; 0.73 (0.54-0.97)</td>
</tr>
<tr>
<td>Czech</td>
<td>Female</td>
<td>834 (309/525)</td>
<td>T 0.34/0.32</td>
<td>0.811</td>
</tr>
<tr>
<td>Swedish</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

MAF, minor allele frequency; T1D, type 1 diabetes; DN, diabetic nephropathy.

Possible genetic interaction between IGF2BP2 rs4402960 and IGF2 rs10770125

IGF2 rs10770125 was associated with DN in male patients with T1D (P=0.037, OR=1.49 95% CI=1.02-2.16) but not in females of the GoKinD population. A possible interaction between IGF2BP2 rs4402960 and IGF2 rs10770125 was observed (P=0.05) by using GMDR program.

Igf2bp2 gene expression at mRNA and protein levels in animal model

Igf2bp2 mRNA expression in kidneys of db/db mice at the age of 5 weeks before onset of DN was increased compared with control mice (P=0.008). However, there was no difference between db/db and control mice at the age of 26 weeks. The results were consistent with protein expression.
5 DISCUSSION AND FUTURE PERSPECTIVE

In this thesis, we have selected the *IGF1, IGFBP1, IGF2* and *IGF2BP2* genes from the IGF-IGFBP axis and analyzed the association of DNA polymorphisms, methylation and serum levels of these molecules in T1D, T2D and DN. The results from our studies provide further genetic and new epigenetic information to understand their roles in diabetes and DN.

5.1 *IGF1* SNP AND DNA METHYLATION

Genetic and epigenetic factors may interact and play a role in the pathogenesis of diabetes. A recent study indicated that *NDUFB6* DNA polymorphisms interacted with methylation changes in regulation of the gene expression in human muscle and increased age-dependent susceptibility to IR [165]. Another study suggested that the T2D associated genetic variants might contribute to the phenotypes of the disease via differential DNA methylation by introducing or removing a CpG site [166]. The SNP rs35767 in the *IGF1* gene was reported to be associated with IR and circulating IGF-I levels [52]. In the study I, we analyzed *IGF1* DNA methylation levels at SNP-CpG site and implied that this SNP rs35767 might regulate IGF-I protein levels by removing a CpG site and changing methylation levels. In addition, the increased DNA methylation at a CpG site in the promoter and decreased circulating IGF-I levels were suggested to be associated with T2D. However, we had no opportunity to analyze *IGF1* mRNA expression in tissue samples, which is a limitation in our study. Further investigation of gene expression at mRNA and protein levels together with analyses of DNA polymorphisms and methylation in the *IGF1* as well as *IGF1R* genes with blood and tissue samples has been taken into our consideration.

5.2 *IGFBP1* DNA METHYLATION ALTERATION AND INSULIN

Data from studies II and III for the first time demonstrated that *IGFBP1* DNA methylation levels were decreased in T1D patients but increased in T2D patients in comparison with NGT subjects. No change of *IGFBP1* DNA methylation levels between T1D patients with and without DN was observed. Both T1D and T2D are characterized by hyperglycemia, but the underlying pathogenic mechanisms may be different. T1D develops on the basis of autoimmune destruction of pancreatic β cells, which usually leads to absolute insulin deficiency. The hyperglycemia in T2D results from a combination of impaired insulin secretion and insulin resistance.

We attempted to investigate the mechanism of *IGFBP1* DNA methylation alteration behind T1D and T2D. In our laboratory, we have stored liver tissues dissected from Goto-Kakizaki (GK) and Wistar rats. GK rat is a hereditary non-obese animal model of T2D, which exhibits a markedly reduced glucose-induced insulin release [167], [168]. Compared with age-matched Wistar rats, GK rats have 10-30% less average body weight but 50-55% higher blood glucose level. In adult GK rats, the β cell mass is usually reduced up to 60% with markedly decreased insulin secretion [167]. IR has been demonstrated in liver, adipose tissue and skeletal muscles of GK rats [169].
We used the tissue samples from 20 male GK rats at the age of 2.5 months. Half of them were implanted with a sustained insulin release chip containing 26±2 mg microcrystallized bovine insulin palmitic acid for 15 days. Ten age-matched male Wistar rats from a local breeder were used as controls. Comparison tests of plasma glucose levels, body weight and serum insulin levels among Wistar, GK rats with and without insulin treatment at sacrifice were done in our previous report [170]. Plasma glucose levels in GK rats (9.6±1.6 mmol/l) were significantly higher compared with Wistar (5.8±0.1 mmol/l, P=0.008) and insulin-treated GK rats (5.3±0.3 mmol/l, P=0.005). Body weights of GK rats with (287.8±3.2 g) and without insulin treatment (273.1±3.3 g) were lower compared with Wistar rats (400.2±8.6 g, P<0.001). Serum insulin levels in GK and Wistar rats were similar (38.4±3.4 vs. 46.0±7.3 µU/ml, P=0.657). As expected, serum insulin levels in insulin-treated GK rats were significantly higher compared with Wistar and GK rats without insulin treatment (68.4±6.0 vs. 46.0±7.3 and 38.4±3.4 µU/ml, P=0.035/0.006).

We extracted DNA samples from liver tissues and then performed the *igfbp1* DNA methylation analysis with the similar protocol as we used in human *IGFBP1* gene studies (Studies II and III). There was no PyroMark assay from Qiagen available for analysis of rat *igfbp1* DNA methylation levels. We thus designed a PyroMark assay for analysis of CpG sites in the putative promoter of the *igfbp1* gene in rat. The preliminary results showed that liver tissue DNA methylation levels of the *igfbp1* gene in GK rats both with or without insulin treatment were significantly increased compared with Wistar rats (39.9% or 41.3% vs. 36.4%, P=0.023 and 0.004, respectively) which are consistent with the data in Study II. But no significant difference of *igfbp1* DNA methylation levels in liver between GK rats with and without insulin treatment was found, although GK rats with insulin treatment had decreased plasma glucose and increased serum insulin levels. This is properly due to the relatively short period (14 days) of insulin treatment or the different sequences of CpG sites of the *IGFBP1* gene between two species i.e. human and rat.

5.3 CIRCULATING IGFBP-1 LEVELS IN T1D, T2D AND DN

Circulating IGFBP-1 derives from liver, and hepatic production of IGFBP-1 is inhibited by portal insulin concentration [124]. In the Study II, T2D patients that we selected into our study were either newly diagnosed or with short duration, because of hyperinsulinemia, serum IGFBP-1 levels in T2D patients were decreased. In the Study III, in consistent with previous report [171], due to the endogenous insulin deficiency, serum levels of IGFBP-1 in T1D patients were increased. We further demonstrated that T1D patients with DN had higher serum IGFBP-1 levels than T1D patients without DN. Kidney accumulation of IGF-I has been shown to be associated with kidney hypertrophy, which is an early feature of DN [137]. The increased IGFBP-1 may allow trapping IGF-I accumulated in kidneys and consequently contribute to the development of DN.

5.4 DNA METHYLATION ANALYSES WITH BLOOD AND TISSUE SAMPLES

DNA from peripheral blood is clinical accessible and easy to handle in laboratory processing, which allows DNA methylation analysis possible in population based
studies. Blood-based DNA methylation has been suggested to be a potential biomarker for cancer molecular epidemiology [172]. However, the source of DNA in peripheral blood is originated from several different cell types, and each of them may exhibit specific methylation pattern [173], [174]. Furthermore, the correlation between methylation at specific loci in peripheral blood DNA and tissue DNA is uncertain. Nevertheless, tissue/cell specific DNA methylation analysis could directly link to gene expression regulation and provide more information for better understand the molecular mechanism behind.

In our studies, we analyzed the blood DNA methylation levels of \textit{IGF1} and \textit{IGFBP1} in diabetes cases and non-diabetic controls. We did not adjust the results according cell-type proportions due to the lack of information. However, it has been reported that there was no significant differences of the portions of sub-cell-types in leucocytes between cases and controls, and adjustment for blood lineage count did not affect the association of methylation with diabetes incidence [175], although high white blood cell count has been suggested to associate with the development of T2D [176]. Furthermore, Dick and his colleagues recently performed a genome-wide analysis to study the association between DNA methylation and BMI. They reported an association between BMI and specific \textit{HIF3A} methylation sites in both blood and adipose tissue DNAs, indicating the assessment of DNA methylation in whole blood is worthwhile for epigenetic study in metabolic diseases [89].

Several major components in the IGF-IGFB axis are predominately expressed in liver and liver tissue specific DNA methylation will determine the hepatic expression of the studied genes. Furthermore, hepatic IR is a feature of T2D. Liver is one of the target tissues for IR. Thus, DNA methylation analysis with liver tissue samples could help us to better understand the correlation between IGFBP-1 DNA methylation and circulating levels as well as the relation with IGF-1 and the pathogenesis of IR in T2D.

In addition, IGFBP-2 is an obesity-driven hormone and secreted from white adipose tissues. The deposit-specific modulation of IGFBP-2 in adipose tissues has been demonstrated in human and mice [177]. Decreased circulating levels of IGFBP-2 are associated with metabolic syndrome such as obesity and diabetes [135], [136]. There is a CpG island located in the promoter and first exon of the \textit{IGFBP2} gene. We have collected abdominal and visceral white adipose tissues from non-diabetic subjects, T2D patients and obese subjects in our laboratory. Tissue specific DNA methylation analysis of \textit{IGFBP2} may help us to understand the deposit-specific modulation of IGFBP-2 and its effect in T2D and obesity.

5.5 GENETIC ASSOCIATION OF IGF2BP2 WITH T2D AND T1D-DN

The importance of the \textit{IGF2BP2} gene in diabetes has been noticed since identification of the intronic SNP in \textit{IGF2BP2} associated with T2D by GWAS. In the Study IV, we successfully replicated the association of SNP rs4402960 in the \textit{IGF2BP2} gene with T2D. Furthermore, we found that this SNP was associated with DN in males and it may interact with \textit{IGF2} SNP rs10770125 in the development of DN. Combined with the previous studies from our group, we identified that several genetic polymorphisms in the \textit{Sox2}, \textit{AdipoQ}, \textit{MME} and \textit{IGF2BP2} genes on chromosome 3q were associated with
DN in T1D. The cumulative genetic effects of all these variants in DN may need to be evaluated in the future.

5.6 OTHER STUDY RELATED TO EPIGENETIC MECHANISM

In this thesis, we have analyzed DNA methylation in diabetes and DN. From our studies, we have gained working experience and knowledge in epigenetics of diabetes and DN. To fully understand the epigenetic effects of the genes in diabetes and DN, we need to extend our research in relation with other epigenetic mechanisms.

DNA methylation and histone modification are two major epigenetic changes. Histone modifications usually include lysine acetylation, lysine/arginine methylation, serine/threonine phosphorylation, and lysine ubiquitination/sumoylation [178]. Histone acetyltransferases (HATs) is associated with gene activation by mediating histone lysine acetylation, whereas histone deacetylases (HDACs) remove lysine acetylation therefore repress gene expression. Both HATs and HDACs have been found to play important roles in diabetes pathogenesis [179]. The recruitment of HDAC1 and deacetylation of histones H3 and H4 can repress the expression of Pdx1 gene, which is a transcription factor critical for β cell function and development [180].

In addition, microRNAs (miRNA) are involved in epigenetic regulation. A miRNA is a non-coding RNA, which suppresses the gene expression at posttranscriptional levels. Several miRNAs have been identified to play physiological roles in tissues in which diabetic complications occur [181]. Wang et al. suggested that up-regulation of miR-320 and consequently decreased expression of IGF-I were likely related to impaired angiogenesis in diabetes [182].

Taking together, there is a need to establish necessary laboratory and collaboration conditions to study other epigenetic mechanisms and their interactions with DNA methylation for better understanding the epigenetic regulation of IGF-IGFBP axis in diabetes and diabetic complications.
6 CONCLUSIONS

- Increased DNA methylation in the promoter of the *IGF1* gene and decreased circulating IGF-I levels were associated with T2D in Swedish males.

- *IGFBP1* DNA methylation levels were decreased in T1D patients but increased in T2D patients in comparison with NGT subjects in Swedish males. Furthermore, decreased and increased IGFBP-1 serum levels were associated with T2D and T1D respectively.

- The *IGF2BP2* polymorphism rs4402960 was associated with T2D. This polymorphism and rs10770125 in the *IGF2* gene were found to be associated with DN in male T1D patients.

In conclusion, our studies provide evidence that the *IGF1, IGF2, IGFBP1* and *IGF2BP2* genes have genetic and epigenetic effects in diabetes and DN. To better understand the importance of our findings, further investigations of tissue specific DNA methylation levels and their impacts on translated proteins are needed. Our findings are summarized in the figure below. It is not excluded that the epigenetic changes precede and/or are involved in the development of diabetes and DN.
ACKNOWLEDGEMENTS

The work of the thesis was performed at Rolf Luft Research Center for Diabetes and Endocrinology, Department of Molecular Medicine and Surgery, Karolinska Institutet. I would like to express my gratitude to everyone who has supported me during these years, particularly to:

My main supervisor Kerstin Brismar, for accepting me to study in KI, for your continuous support and excellent guidance, for sharing your broad knowledge and abundant working experience of clinical and basic research in metabolism and endocrinology, particularly in the IGF system. Your enthusiasm for science really inspires me. You are always very kind and positive, and it is really enjoyable to work under your supervision.

My co-supervisor Harvest F Gu, for introducing me to the fields of genetics and epigenetics, for always being available for scientific discussion, for teaching me everything in the lab, for patiently educating me from a medical student with limited knowledge on molecular biology to a junior researcher. Thank you for letting me to be aware of the importance of working independently and collaboratively.

Ying Liu at Peking Union Medical College, for your encouragement as my external mentor.

Claes-Göran Östenson, for valuable discussion and comments on my studies and kindly involving me in your group activities.

Kerstin Hall, for your valuable and constructive comments in Friday meetings and my half-time seminar.

Gustav Dallner, for inspiring me with your positive, optimistic and enthusiastic attitudes towards science.

Sergiu-Bogdan Catrina, for organizing the breakfast seminar and constructive comments during our discussion in lab meetings.

My co-authors in our studies, Agneta Hilding, for the great help in statistical analyses. Henrik Falhammar, Eva Horova, Anna Möllsten and Martin Prazny, for excellent clinical data collection and nice collaboration, particularly to Henrik, for valuable and long discussion on clinical data. Tomas Ekström and Louise Sjöholm for the introduction and discussion on epigenetics.

Monika Pettersson for valuable discussion and assistance in assay design, Elvi Sandberg and Inga-Lena Wivall, for excellent laboratory assistance.

Thanks to all my friends at M1:03: Anna Kistner, Anneli Björklund, Carole Muller, Elizabeth Noren-Krog, Faradianna Lokman, Galyna Bryzgalova, Julien Pelletier,
Mohammed Seed Ahmed, Neil Portwood, Norhashimah Abu Seman, Saad Alqahtani, Senthil Vasan, Silvia Zambrana, Tina Wirström and Zuheng Ma, for creating an enjoyable working environment, for sharing bits and pieces of life. We are such an international and harmonious family! Keep in touch no matter where we will be in the future.

To all the colleagues at L1: 01: Christina Bark, Elisabete Forsberg, Ileana Botusan, Ishrath Ansurudeen, Ismael Valladolid, Jacob Grunler, Jing Wang, Marianna delsole, Michael Tekle, Noah Moruzzi, Teresa Daräio, Vivek Sunkari and Xiaowei Zheng for our scientific discussion in group meeting every week, for great time we had together.

I would like to thank the administrative staff at MMK: Ann-Brit Wikström, Britt-Marie Witasp, Kerstin Florell and Katarina Breitholtz for all the kind help, Jan-Erik Karre, Lennart Helleday and Thomas Westerberg for professional IT support.

To all my Chinese friends in Stockholm: Bin Li, Bojing Liu, Chengjun Sun, Chen Suo, Ci Song, Dong Yang, Ge Wu, Huan Song, Hongchang Shen, Hongya Han, Jia Cao, Jianren Song, Jia Sun, Yan, Jianwei Zhu, Jiaqi Huang, Jiangnan Luo, Kai Du, Lidi Xu, Meiqiongzi Zhang, Meng Xu, Na Guan, Peng Zhang, Qiang Zhang, Qin Xiao, Rui Wang, Simei Yu, Tiantian Liu, Ting Jia, Tingting Lin, Xiaoyan Huang, Xiaozhen Li, Xinmin Wang, Yixin Wang, Yang Xuan, Ying Qu, Yuning Zhang and Zheng Chang. Thank you all for the good time we have shared, for all the trips, parties, games, laughter and all the fun we had together, for making my life here not feeling lonely and more colorful. The past five years and all the moments with you will be a great memory in my life. Cheers for our friendship!

Most important, thanks to my entire family, for every video chats during weekends, for every family gathering when I was back for vacation, for your support and endless love. I will soon be back to be a part of your life.

My aunt in Stockholm, thank you for taking care of me as a member of your family, Tianlin and Karolin, thank you for the good time we had together, for all the communications about life and future.

My girlfriend Wei Dong, for your constant love and faith in me, for your understanding and support.

I owe my biggest thanks to my dearest parents, who never stop giving of themselves in countless ways and encourage me at every step of my life. I am sorry I was not by your side when you really need me. Thank you for your understanding, support, encouragement and all positive attitudes when I study abroad.

This study was supported by Family Erling-Persson foundation, Swedish Research Council, Swedish Diabetes Association, Stig and Gunborg Westmans foundation and funds from Karolinska Institutet.
8 REFERENCES


“DNA methylation and body-mass index: a genome-wide analysis,” Lancet, Mar. 13 AD.


